Records with non-Latin scripts in Aleph

All non-Latin scripts supported by OCLC may be used in Aleph. Furthermore, non-Latin scripts not supported by OCLC may also be used in Aleph, though OCLC may not retain the vernacular portion of the record in their database.

Please see OCLC's documentation on "International Cataloging" for a current list of scripts that they support.

When importing records from Connexion, your client should be configured to export records in UTF-8 (see Connexion Tools menu, Export tab, Record Characteristics button).

Paired fields

Non-Latin data in Aleph is stored in paired fields. The first field of a pair should contain the romanized text, with the second field containing the equivalent non-Latin script. (Note that this is opposite of how the record appears in OCLC Connexion, which displays the non-Latin field first). The paired fields are connected through the use of a subfield 6:

```
245 00 6 01
   a Ba zheng ma nian :
   b Guo li gu gong bo wu yuan ba shi nian de dian di huai xiang /
   c [zhu bian Cai Meifen].
245 00 6 01
   a 八十年念 :
   b 国立故宫博物院八十年的点滴怀想 /
   c [主编蔡敏芬].
250 ___ 6 02
   a Chu ban.
250 ___ 6 02
   a 初版.
260 ___ 6 03
   a Taibei Shi :
   b Guo li gu gong bo wu yuan,
   c 2006.
260 ___ 6 03
   a 台北市 :
   b 国立故宫博物院,
   c 2006.
```

Subfield 6 should always be the first subfield present in a paired tag. It should be a two-digit number, with numbers less than 10 beginning with a zero. Subfield 6 should be unique for each set of paired fields. Improper coding of paired fields will result in errors when the records are sent to OCLC.

More than one non-Latin script may be present in a single record:
If a tag containing non-Latin script does not have a transliterated pair, such as a contents note (tag 505), staff should use a subfield 6 containing 00.

For information on paired fields in OCLC, please see OCLC's "International Cataloging" document.

Notes

Staff. Staff editing non-Latin data fields and their romanized counterparts should have a strong understanding of the language and the ALA / Library of Congress romanization scheme for that language, as well as the conventions for inputting paired fields in the MARC 21 Bibliographic Format and the way these are applied in the HOLLIS catalog. Staff may input or update non-Latin fields in bibliographic records using the appropriate Windows Keyboard/IME, which can be set up using the language and regional options available in the Windows Control Panel.

Indexes. All non-Latin scripts are indexed in the main keyword indexes. CJK scripts have their own browse indexes. Other non-Latin scripts file in the main browse indexes based on the Unicode values of the individual characters.

Filing indicators. Filing indicators for titles in paired fields may differ, depending on the language.

LCSH source indicator. Staff should use the appropriate source indicator in tag 6XX when inputting non-Latin terms. Usually this will be 4 ("Source not specified"), because the LC authority record does not contain an authoritative form for an LC heading in the non-Latin form. CONSER policies may differ.

Tag 066. This tag may be present in records which were imported in the MARC8 character encoding scheme. For records imported from OCLC in UTF-8 encoding (which began at Harvard in August 2008), this tag will not be present.

Comments / Questions

Please contact the Aleph Support Center with any questions or comments:
http://nrs.harvard.edu/urn-3:hul.ois:alephhelpcharset